AI and Personalization

Omid Rafieian* — Cornell Tech and Cornell University

Hema Yoganarasimhan* — University of Washington

Abstract

This paper reviews the recent developments at the intersection of personalization and AI in marketing and related fields. We provide a formal definition of personalized policy and review the methodological approaches available for personalization. We discuss scalability, generalizability, and counterfactual validity issues and briefly touch upon advanced methods for online/interactive/dynamic settings. We then summarize the three evaluation approaches for static policies – the Direct method, the Inverse Propensity Score estimator, and the Doubly Robust method. Next, we present a summary of the evaluation approaches for special cases such as continuous actions and dynamic settings. We then summarize the findings on the returns to personalization across various domains, including content recommendation, advertising, and promotions. Next, we discuss the work on the intersection between personalization and welfare. We focus on four of these welfare notions that have been studied in the literature: (1) search costs, (2) privacy, (3) fairness, and (4) polarization. We conclude with a discussion of the remaining challenges and some directions for future research.

*Please address all correspondence to: or83@cornell.edu, hemay@uw.edu.

1 Introduction

Effective personalization at scale remains the ultimate holy grail for marketers and has drawn interest from both academics and practitioners for a long period of time. Nevertheless, until recently, individual-level personalization at scale was, at best, a theoretical possibility. With advances in computing power and data storage, combined with the theoretical developments at the intersection of machine learning and causal inference, we are now closer than ever to achieving this goal. A growing body of research over the last few years, spanning marketing, computer science, economics, and statistics, has collectively attacked this problem from various angles and provided us with solutions that are both theoretically well-founded and practically feasible. Today, these solutions and ideas have been adopted by firms across various business settings, from news and content personalization on websites to the personalization of promotions by retailers. These applications, in turn, have opened our eyes to yet unsolved challenges related to this problem. Further, they have enabled us to quantify the revenue and business impact of personalization and better understand the welfare and public policy implications of user-level personalization.

This chapter reviews the recent developments at the intersection of machine learning, artificial intelligence, and personalization. In §2, we first present a mathematical definition of the personalization problem. This definition will form the foundation on which all the solution concepts will be built. Next, in §3, we discuss the methods available for personalized policy design. We also explain the pros and cons of each and discuss how we need to combine ideas from causal inference and machine learning to effectively design policies that have counterfactual validity. In §4, we present the approaches available to evaluate personalized policies and cover some best practices. Then, in §5, we discuss the substantive findings on personalization in a variety of settings. In particular, we highlight the methods and domains that have seen successful applications of personalized models in both academic research and practice. An interesting aspect of personalization is that it is closely linked to many different aspects of consumer and social welfare, such as fairness, privacy, and polarization. In §6, we discuss these ideas and examine the personalization problem through these lenses. Finally, §7, we conclude with a discussion of the remaining challenges, unsolved problems, and some directions for future research.

An important caveat before we proceed: The topic of personalization, by definition, has a large scope and has been the focus of research across many disciplines. Even within marketing, personalization has been studied for many years from different angles (e.g., game-theoretic models, substantive empirical work, and consumer behavior studies). In this review, we restrict our focus to methods and studies that fall at the intersection of machine learning and personalization. We mainly discuss the recently developed machine learning methods on this topic and the papers that employ these methods. We refer interested readers to Murthi and Sarkar (2003), Steckel et al. (2005), Proserpio et al. (2020), and Liaukonyte (2021) for earlier reviews on personalization from other perspectives.

2 Problem Definition

We start with a formal definition of personalization. Broadly, we can define personalization as the opposite of the case where a firm implements a uniform policy for the entire population. For example, a uniform policy would offer everyone the same discount rate in a promotional context. In contrast, a personalized discount policy would provide different discount rates to consumers based on observed features of the consumer. The granularity of personalization can vary, from broad segments (based on demographics) to highly individualized strategies where each consumer gets a treatment based on a complex model trained on a high-dimensional set of user-level observables.

Therefore, firms' ability to personalize is highly dependent on their ability to differentiate between individuals. We assume that firms can differentiate between individuals based on the feature vector X_i, which denotes the characteristics associated with the user/observation i. This can include any individual-level features, such as demographics or the behavioral history of the individual (which can consist of users' responses to past marketing actions by the firm). Further, let the firm's action set $\mathcal{W}$ denote the set of actions that the firm can choose from and let Y_i be the outcome of interest that the firm seeks to optimize. Then the firm's goal is to select an action W_i for each user i such that it maximizes the expected outcome for that user, i.e., $E(Y_i|X_i, W_i)$. For example, in the promotion example above, the discount amount can be interpreted as the action, and the dollar value of the purchase can be the outcome of interest for the firm. We now use this notation to define a personalized policy or strategy as follows:

Definition 1. *Let $\mathcal{X}$ denote the set of all the features associated with user characteristics. A personalized policy π is defined as a mapping from the set of characteristics $\mathcal{X}$ to the set of actions $\mathcal{W}$, such that each $X_i \in \mathcal{X}$ is assigned to only one action, i.e., $\pi : \mathcal{X} \to \mathcal{W}$.*

This definition highlights the difference between a personalized policy and a uniform policy. In a uniform policy, the policy function π only has one output for the entire population, i.e., $\pi(X_i) = \pi(X_j)\ \forall\ i$ and j. In contrast, the goal of a personalized policy is to find the right action for each observation to maximize a certain objective, which is often characterized by the outcome of interest Y_i. To the extent that there is individual-level heterogeneity in the users' response to actions and the supervised learning model used can capture this personalized policy will differ from a uniform policy. However, identifying heterogeneity at the individual level is often challenging. While the space of uniform policies is relatively small, the space of personalized policies is very large. Therefore, learning a personalized policy requires large-scale data sets. In fact, the narrower we want to target, the larger the scale of data needs to be. Further, to effectively develop personalized policies from data, we need flexible methods to learn complex patterns from the data. In the next section, we review methods that can overcome these challenges and personalize effectively.

3 Methodological Approaches to Personalization

In this section, we present an overview of the methods and approaches used to derive a personalized policy, as described in Definition 1.

In one of the earliest empirical papers on personalization, Rossi et al. (1996) develop a Bayesian framework for coupon personalization in a grocery context. They use counterfactual simulations to document the possibility of substantial gains when coupons are targeted based on customers' past purchase history. In building their framework, Rossi et al. (1996) noted the "continued decline in information processing and storage costs" as the main reason why we can target marketing activities at the individual level. Indeed, with the growth of Internet firms and the increased ability to deliver personalized policies, led to a body of work that focused on building personalization frameworks in other domains, such as movie recommendations (Ansari et al., 2000; Ying et al., 2006; Chung and Rao, 2012), e-mail personalization (Ansari and Mela, 2003), personalized web design through browsing behavior (Montgomery et al., 2004), promotion customization (Zhang and Krishnamurthi, 2004), and music recommendation (Chung et al., 2009). At a fundamental level, these approaches focused on outcome modeling and often leveraged the hierarchical Bayes model or flexible regressions to capture user-level heterogeneity.

This early empirical work on personalization, in conjunction with the practical industry-level challenges, opened up avenues for further methodological research that are: (1) scalable, (2) generalizable and robust to confounding, (3) adaptive in real-time, and (4) dynamic or forward-looking. In the following sections, we discuss the methodological approaches along these four dimensions.

3.1 Scalability

Greater personalization often requires more data. Naturally, it is impossible to deliver an effective fine-grained personalized policy with limited data. Further, most platforms that use personalization collect data at a massive scale. As a result, personalization methods need to be scalable. This section focuses on two broad sets of approaches proposed in the literature: (1) dimensionality reduction methods and (2) scalable supervised learning algorithms.

The first set of papers focuses on reducing the dimensionality of the problem. The canonical example of such an approach is the well-known collaborative filtering problem, where the problem in Definition 1 is formulated by a preference matrix where users are represented by rows and actions by columns, and the value is the outcome of interest (Goldberg et al., 1992; Sarwar et al., 2001). While these methods represent each user in a row, they do not use the attributes X_i. The explicit structural assumption in these models is the low-rank structure of the preference matrix. As such, the preference for each user is represented as a composite of factors in this low-rank representation. Intuitively, this approach uses the similarities in user preferences to complete the preference matrix. Low-rank methods have been widely used in recommender systems in the industry (Linden et al., 2003) and motivated a series of theoretical papers given their interesting statistical properties (Candès and Recht, 2009; Candès and Tao, 2010; Recht, 2011). Recent work in this domain extends this work to more specific application domains, when side information is available (Farias and Li, 2019), dynamic assortment personalization where online exploration is necessary (Kallus and Udell, 2020), and cases where online exploration can result in customer disengagement (Bastani et al., 2021b).

Another common approach to dimensionality reduction is to use probabilistic topic models that map a high-dimensional space to a low-dimensional set of topics (Blei et al., 2003). While the typical application of topic models is in text mining, marketing scholars have also applied these methods to purchase history data to predict consumers' behavior, i.e., what would they buy next, and thus achieving greater personalization ability (Jacobs et al., 2016, 2021). Recent work in this domain combines probabilistic topic models with matrix factorization approaches to develop scalable personalization algorithms (Ansari et al., 2018). A key advantage of these newer approaches is that, unlike most common matrix factorization models, these approaches take all the information in X_i into account.

The second set of papers in this domain addresses the scalability problem using flexible supervised learning algorithms, such as Deep Neural Networks, XGBoost, and Random Forests.[1] As such, these methods feed the high-dimensional set of characteristics X_i and actions W_i as inputs to a supervised learning algorithm to learn a flexible function f that estimates the outcome Y_i with high predictive accuracy. These approaches are increasingly more feasible with the advances in high-performance computing. This has motivated practitioners to move to this kind of approach for a variety of problems, such as CTR prediction problems for personalized ad recommendation (McMahan et al., 2013; He et al., 2014; Chapelle et al., 2015), movie recommendation (Li et al., 2015), and news recommendation (Wang et al., 2017a). Please see Singhal et al. (2017) for a summary of the use of deep learning in modern recommender systems. One of the key reasons

[1]Another set of scalable approaches include Bayesian methods such as distributed MCMC and stochastic variational inference. These methods are generally used when quantifying uncertainty in our predictions is a top priority. Please see Angelino et al. (2016) for an excellent summary of this stream of methodological work.

for the popularity of these methods is their ability to leverage high-dimensional data, such as the visual content of ads or the text of news articles, for the task of prediction. While these approaches avoid information loss by avoiding low-rank assumptions, they often lack theoretical guarantees on generalizability. In particular, in contexts without enough randomization, these models give sub-optimal policies due to the confoundedness of actions seen in the data.

3.2 Generalizability and Counterfactual Validity

While supervised learning algorithms can learn complex patterns from data, they are only generalizable to instances with the same joint distribution as the data used for training them. Suppose an action-covariate pair (W_i–X_i) had zero probability of being observed in the training data. In that case, there is no guarantee that these methods provide an accurate estimate of the outcome for that action-covariate pair. This is particularly problematic because developing a personalized policy requires evaluating all possible actions for any given X_i and selecting the action that maximizes some outcome interest (in expectation). Further, the primary purpose of a personalized policy is to change the current policy, thereby shifting the current data distribution. Thus, personalization algorithms need to predict counterfactual scenarios accurately.

This challenge motivated a growing body of work that brought insights from the causal inference literature to increase the counterfactual validity of personalization algorithms. The fundamental requirement in this stream of work is a degree of randomization in actions implemented in the training data that satisfies two key assumptions: (1) *overlap*, and (2) *unconfoundedness*. Overlap assumption states that we can consistently predict the outcome for an instance if and only if the action had a non-zero propensity score of occurring in that instance in the data. Unconfoundedness assumption requires the action to be exogenously chosen conditional on observed variables. Research in this domain has used three different approaches to increase the counterfactual validity of the personalization algorithm:

- One approach is to use outcome modeling in fully randomized environments (Hill, 2011; Ascarza, 2018), or with explicit controls for potential confounding factors and propensity scores (Rafieian and Yoganarasimhan, 2021). This approach is akin to early work on personalization and supervised learning methods that aim to estimate the outcome through a function, i.e., $\hat{Y}_i = f(X_i, W_i)$. Note that any supervised learning model can be used to learn $f(X, W)$ – it can be something as simple as a linear regression or an ensemble of different state-of-the-art machine learning methods. As long as the function f can output an action based on X_i, we can use it generate a personalized policy; usually by picking the policy that maximizes expected reward $R(\pi, Y) = \frac{1}{N}\left[\sum_{i=1}^{N} Y(X_i, W_i^{\pi})\right]$.
- The second approach is to directly learn the optimal personalized policy from data using the

propensity scores (Swaminathan and Joachims, 2015; Kitagawa and Tetenov, 2018; Kallus and Zhou, 2021a; Athey and Wager, 2021). These approaches often develop an unbiased estimator of the outcome under a policy and search in the full space of policies to find the optimal policy.

- The third approach is to identify the causal effect of any given action for each instance in the data through estimating the *conditional average treatment effect (CATE)*, using different machine learning methods, such as decision trees (Rzepakowski and Jaroszewicz, 2012; Athey and Imbens, 2016), random forests (Guelman et al., 2015; Wager and Athey, 2018; Athey et al., 2019), deep learning (Shalit et al., 2017). Some of the papers in this domain provide a framework that works for any generic machine learning method (Chernozhukov et al., 2018; Hitsch and Misra, 2018; Nie and Wager, 2021). The fundamental idea in this literature is to use two techniques: (1) orthogonalization to exploit the randomization in the assignment to actions and make estimation robust to confounding, and (2) cross-fitting to prevent machine learning methods from overfitting.

3.3 Online and Interactive Methods

As discussed earlier, personalization algorithms generally require large-scale data. Further, platforms that implement personalized policies often face a fast-paced environment, where the set of actions changes regularly (e.g., entry of new ads, a daily update of news articles). These challenges led to the development of interactive frameworks that actively explore the action space and efficiently find the optimal personalized policy by finding the right balance between exploration and exploitation. To capture the exploration-exploitation trade-off, we can model this problem as a multi-armed bandit with information about the context through the set of characteristics X_i that allows the algorithm to make personalized decisions given the context (Auer, 2002). This problem is referred to as "contextual bandit," and the goal is often to minimize some notion of regret over time (Langford and Zhang, 2007): if the true optimal (oracle) personalized policy is π^* and the contextual bandit policy is π, we can define regret as the difference between reward under the true optimal policy and the contextual bandit policy.

Contextual bandits have been applied in a variety of domains such as news personalization (Li et al., 2010), website and banner advertising morphing based on consumers' cognitive styles (Hauser et al., 2009; Urban et al., 2014), mobile health (Tewari and Murphy, 2017), and content recommendation (Agarwal et al., 2016). These papers aim to develop algorithms for the specific problem at hand and establish the empirical gains from using these algorithms. Another stream of work in the domain of contextual bandits has mainly focused on the theoretical aspects of the problem, such as regret bounds, complexity, and exploration (Foster and Rakhlin, 2020; Bastani and Bayati, 2020; Simchi-Levi and Xu, 2021; Bastani et al., 2021a). Despite this growth on practical and theoretical fronts, many questions remain unanswered in this literature, and there are many avenues for future research. Please see Bietti et al. (2021) for an excellent summary of the recent work in this literature and the remaining challenges.

3.4 Dynamic Methods

The vast majority of personalization algorithms are designed for static cases with a greedy objective function, where the firm only focuses on immediate rewards. This has led to the criticism that personalization algorithms only help achieve short-term rewards. To the extent that the short- and long-term rewards conflict, static personalization algorithms can generate sub-optimal long-term outcomes. For example, personalized policies can increase immediate engagement but result in higher churn among the users on a platform. To address this issue, an emerging stream of work has focused on developing a forward-looking objective for decision-making that accounts for the interdependence between the actions over time. This is usually done by formulating the problem as a Markov Decision Process that captures both immediate rewards and expected future rewards.

A series of papers have taken this forward-looking approach and designed dynamic personalized policies to optimize long-term value in a variety of setting, e.g., personalized ad recommendations (Theocharous et al., 2015), optimizing user engagement by sequencing ads and capturing dynamic ad effects (Rafieian, 2022), website design (Liberali and Ferecatu, 2022), investment strategies in robo-advising (Capponi et al., 2021), long-term impact of mobile health interventions (Liao et al., 2021), and educational outcomes in games (Mandel et al., 2014).

4 Evaluation

Once we develop a personalized policy, we need to evaluate this policy to assess its performance before adopting it. Evaluation is a critical step in any study on personalization.

There are two main approaches to evaluation: (1) field experimentation and (2) counterfactual policy evaluation. Field experimentation is the ideal gold standard, and it allows us to test the personalized policy against a series of benchmark policies and estimate the total rewards and average treatment effect (ATE) of the personalized policy. This is generally what happens before a personalized policy gets implemented in online platforms through A/B testing (Kohavi et al., 2020). However, field experiments are not always feasible because designing and running a large-scale experiment is costly in time and money. Therefore, we need methods that reliably evaluate the performance of a personalized policy based on existing data before considering a field test.[2] This is where counterfactual policy evaluation comes in – we want to answer what would happen if the

[2]Further, in some settings, even a field experiment will not give us consistent estimates of the gains from a specific policy; see Goli et al. (2022) for example.

Outcome Y_i	Covariates X_i	Observed Action W_i		Prescribed Action $\pi(X_i)$	Personalized Policy Implemented $Z_i = \mathbf{1}(W_i = \pi(X_i))$
...	...	Action 1	✗	Action 2	0
...	...	Action 3	✗	Action 1	0
...	...	Action 1	⟷	Action 1	1
...	...	Action 2	✗	Action 3	0
...	...	Action 4	⟷	Action 4	1

Figure 1: An overview of matching personalized policy with the one observed in the data. Green arrows indicate a case where the actions match between data and the personalized policy, and red crosses indicate otherwise.

personalized policy were to be adopted.[3]

We review methods for counterfactual policy evaluation that exploit randomization in the actions observed in the data and build on the statistical properties of a randomized field experiment to consistently estimate total rewards and ATE under counterfactual regimes. The key intuition behind these methods is that under some level of randomization, the personalized policy is implemented in some observations in the data just due to randomness. As such, like the fully randomized experiment, some observations received the personalized policy, and some did not. The difference with a fully randomized experiment is that the assignment is not necessarily exogenous. Therefore, if we have the propensity scores for the actions observed in the data, we can consistently estimate the reward under the personalized policy and the ATE. Figure 1 illustrates this point clearly. In this section, we first review three such methods to evaluate counterfactual policies as laid out by Dudík et al. (2011), and then discuss alternatives used in the literature and the challenges therein.

4.1 Direct Method

This method estimates the reward from a proposed policy by predicting the expected outcome and directly plugging in the predicted values. This can be formally written as:

$$\hat{R}_{DM}(\pi, Y) = \frac{1}{N} \sum_{i=1}^{N} \hat{Y}(X_i, \pi(X_i)), \tag{1}$$

where $\hat{Y}(X_i, \pi(X_i))$ is our model-based prediction of the outcome for observation i when she is given the policy-prescribed action $\pi(X_i)$. This method is known as the Direct Method since it directly uses model predictions in the evaluation formula.

The performance of this estimator depends on the predictive accuracy and counterfactual validity

[3]This problem is also referred to as off-policy policy evaluation in the computer science literature.

of $\hat{Y}$. In general, this method will work well if: (1) the model used to predict $\hat{Y}$ was sufficiently flexible (e.g., XGBoost, deep learning methods), and (2) $\hat{Y}$ was estimated on a data set with sufficient randomization that satisfies the overlap and unconfoundedness assumptions. The second condition is generally satisfied if the propensity scores for all actions prescribed by the policy are non-zero and known. However, suppose an action prescribed by the personalized policy has zero propensity score (i.e., it could have never been implemented in the data). In that case, there is no guarantee that a predictive model can accurately estimate this action's outcome. Thus, settings where $\hat{Y}$ was learnt using flexible machine learning models with inherent randomization in actions are ideal use-cases for this estimator. Rafieian and Yoganarasimhan (2021), whose setting satisfies these two requirements perform an empirical exercise where they compare the estimated gains from a model-based approach and show that it is very similar to the estimates from a model-free IPS approach (see below).

4.2 Inverse Propensity Score Estimator

In contrast to the direct method, whose performance is based on the model, the Inverse Propensity Score (IPS) estimator is a *model-free* approach to evaluation. It is based on the idea of importance sampling by Horvitz and Thompson (1952), and is now increasingly used in marketing papers on personalized policy evaluation (Simester et al., 2020a; Rafieian and Yoganarasimhan, 2021; Yoganarasimhan et al., 2022; Goli et al., 2021). For any given policy π, this estimator takes all the observations where the user received the policy-prescribed treatment and scales them up by their propensity to receive the treatment. This gives us a pseudo-population that received the policy-prescribed treatment. Thus, the average outcome for this pseudo-population is an unbiased estimate of the reward for the full population if we were to implement the proposed policy in the field. Formally:

$$\hat{R}_{IPS}(\pi, Y) = \frac{1}{N} \sum_{i=1}^{N} \frac{1[W_i = \pi(X_i)]Y_i}{\hat{e}(W_i)}, \tag{2}$$

where Y_i is the observed outcome for observation i, and $\hat{e}(W_i)$ is the propensity score for action W_i to be implemented in observation i. Again, in settings where the propensities are non-zero and known for the prescribed policy, we can directly use propensity scores to evaluate the performance. The main advantage of this approach over the direct method is that it does not require any predictive model to evaluate the outcome. Instead, it uses actual outcomes observed in the data. However, this approach works well only if propensity scores $\hat{e}(W_i)$ are correct. Further, even if propensities are known but small, the variance of the IPS estimator can be high, which can preclude us from drawing any reliable inference on the performance of comparable policies.

4.3 Doubly Robust Method

The two approaches described earlier require either outcome estimates or propensity scores to be accurate. The problem is, in many cases, we do not which one would be accurate ex-ante. This is where the Doubly Robust (DR) estimator comes in by combining the strengths of both the direct method and the IPS method (Dudík et al., 2011) and guaranteeing consistency and unbiasedness if *either* outcome estimates *or* propensity scores are accurate. DR estimator is defined as follows:

$$\hat{R}_{DR}(\pi, Y) = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{1[W_i = \pi(X_i)](Y_i - \hat{Y}(X_i, \pi(X_i))}{\hat{e}(W_i)} + \hat{Y}(X_i, \pi(X_i))\right], \tag{3}$$

This estimator is ideal in most situations. It is especially powerful in observational settings where propensities are not given (and need to be estimated from data, and hence can be noisy). That said, like the direct method and IPS estimator, if the setting fails to achieve overlap or unconfoundedness assumptions, there is no guarantee that the DR method works. In these settings, the requirement of either accurate outcome estimation or accurate propensity score estimation fails. More recent work has aimed at improving these estimators; see Wang et al. (2017b) as an example.

4.4 Extensions to Special Settings

Much of the discussion in earlier sections primarily focused on a case where there is a discrete and finite action set, and the goal is to deliver a single personalized action at a time in a static setting. However, many cases depart from this setting, imposing challenges to the estimators we discussed. We now discuss these extensions:

- **Continuous actions:** When the action set is continuous (e.g., pricing), it is not clear how IPS and DR methods can be applied since it is not possible to achieve non-zero propensities over the entire action space. For these cases, we can still use the direct method in Equation 1, when there is randomization in actions. Recent work extend doubly robust evaluation methods to cases with continuous treatments (Chernozhukov et al., 2016; Athey and Wager, 2021).
- **Dynamic settings:** Another departure from cases that we already considered is when we want to evaluate a dynamic policy as defined in §3.4. In these cases, propensities are defined not for a single action but for a trajectory of actions. The extension of the direct method is straightforward: we only need to use Equation 1 such that it simulates state transitions (Rafieian, 2022). The extension of both IPS and doubly robust have also been studied in the recent literature on reinforcement learning (Theocharous et al., 2015; Jiang and Li, 2016; Thomas and Brunskill, 2016). The main practical challenge in this approach is that propensity scores capture the propensity for the entire trajectory, which is the product of propensity scores for a single action.

Hence, the overall propensity for a trajectory can become very small, making propensity-based approaches impractical. Please see Kallus and Uehara (2020) for a complete survey of different evaluation methods and a solution under some Markovian simplification.

4.5 Alternative Approaches

We now discuss alternative approaches used for policy evaluation, as well as the challenges and problems involved in using them:

- **Metric-based evaluation:** Since outcome prediction is part of developing personalized policies, some approaches rely on standard goodness-of-fit approaches that measure the out-of-sample predictive accuracy of the model, such as Mean Average Error (MAE), Area Under the Curve (AUC), Entropy, Relative Information Gain (RIG) (Herlocker et al., 2004). While a high predictive accuracy is often necessary for developing an optimal personalized policy, it does not guarantee higher gains from personalization. For example, suppose that a variable like *age* fully determines consumers' response to ads. Even if everyone prefers ad A to ad B, we will have high predictive accuracy because age helps predict the outcome. However, there is no gain from personalization in this case. Yoganarasimhan et al. (2022) empirically demonstrate that policies based on models with lower predictive accuracy often do better than policies based on more accurate models and discuss the reasons for such patterns. Overall, it is essential not to confuse evaluation in terms of predictive accuracy with counterfactual policy evaluation. Most researchers have moved away from metric-based evaluation to counterfactual policy evaluation based on the methods discussed earlier.
- **Structural models:** The problem of counterfactual policy evaluation has long been studied in marketing and economics. The traditional approach to this problem is to build a structural model to estimate a set of model primitives. These primitives can then be used to simulate data under a counterfactual regime. This approach has been used in different contexts in the marketing literature to evaluate the performance of personalized policies (Yao and Mela, 2011; Jiang et al., 2021; Morozov et al., 2021). However, it is important to notice that the nature of identification is not necessarily based on the randomization in actions but rather on the assumptions made on the data-generating process by the structural model. To the extent that these assumptions are correct, the counterfactual policy evaluation delivered by structural models will be valid. However, if these assumptions are incorrect, there is no theoretical guarantee that this approach consistently estimates the rewards under the personalized policy. Thus, these methods should only be used when there is no randomization in actions in the data, and the structural assumptions are mostly harmless.

In sum, we have various methods available for evaluation, and researchers must perform a comprehensive offline evaluation of any personalized policy before considering field implementation. Further, even after the policy has been implemented on the entire population, it is important to compare the performance of the personalized policy against the right control group. For example, Farahat and Bailey (2012) run a targeted advertising campaign and show that if we do not control for the selection bias in the targeted group (i.e., targeted users are more likely to purchase), then simply comparing the response of the targeted users and the untargeted users can significantly overestimate the effects of targeting.

5 Returns to Personalization

One of the major success stories of personalization is the Netflix recommendation algorithm. In 2009-2010, Netflix launched a million-dollar challenge where they asked amateurs to increase the accuracy of their in-house algorithm by 10%. This led to significant interest and investment in personalized recommendations. The Netflix personalized recommendation system is reputed to save the firm over one billion dollars a year and is cited as one of the main reasons for the low churn rate at Netflix (Gomez-Uribe and Hunt, 2015; Kasula, 2020). Similar product recommendation systems are now commonly used by many other content providers such as YouTube, Hulu, Prime Video, and Disney. In this section, we go beyond the industry reports and review the findings of the academic work on the returns to personalization across a variety of domains.

The returns from personalization depend both the action that was personalized (W, which can represent marketing actions such pricing, or promotion) as well as the reward or the outcome of interest that was optimized (Y, which can represent clicks, user engagement, subscription, profit, revenue). The overall effectiveness of any personalized policy thus depends on the extent to which we can move Y by optimizing W as a function of X. In the rest of this section, we categorize the discussion based on the marketing action that was personalized (and provide details on the outcome that was optimized in the respective studies).

- *Content recommendation* is an area where many platforms employ personalized policies, inspired by the Netflix example highlighted above. In this domain, Li et al. (2010) proposed a contextual bandit framework for personalize news recommendation. The explore-exploit paradigm provides a natural solution to the cold start problem in this setting because new articles, whose valuation and fit we need to learn about, are added regularly, while older articles need less exploration. Based on field experiments and deployment on Yahoo!, Li et al. (2010) document a 10% lift in CTR using their contextual bandit model, compared to a naive ϵ-greedy policy. Similarly, experiments at MSN showed a 25-30% increase in CTR with news personalization compared to the editorial ordering (Agarwal et al., 2016). As a result, contextual bandits are now used by other leading news organizations (e.g., New York Times) for news personalization (Coenen, 2019).

- *Advertising* is another area that motivated a vast body of work on personalization (He et al., 2014; McMahan et al., 2013; Rafieian and Yoganarasimhan, 2021; Rafieian, 2022). In particular, Rafieian and Yoganarasimhan (2021) used both the direct method and IPS approach to counterfactual policy evaluation in a context where both overlap and unconfoundedness assumptions are satisfied and documented a 65-67% increase in CTR over the ad allocation policy under a quasi-proportional auction used by the in-app advertising platform. Rafieian (2022) studied a dynamic personalization policy by sequencing ads and showed an over 80% improvement in the total number of clicks over the baseline that randomly allocates ads.
- *Pricing* is another domain where researchers documented positive returns to personalization. Kallus and Zhou (2021b) present an excellent characterization of the different types of price personalization and provide well-grounded notions of fairness and welfare under personalized pricing. They consider two empirical case studies in the domains of elective vaccine and micro-finance lending. They show that personalized pricing (based on covariates) can increase both profits and parity among different sub-groups because personalization can increase take-up. A recent study by Dubé and Misra (2021) echoes these findings (using data from Ziprecruiter). Further, recent work suggests that even when firms cannot personalize prices directly, they can personalize *implicit* prices. For example, using data from large-scale field experiments at Pandora, Goli et al. (2021) shows that ad-supported platforms can increase subscription revenues by personalizing the ad load received by users (i.e., the number of ads seen by the user and their intrusiveness). Extending this research to account for the additional challenges specific to price personalization would be an excellent next step, e.g., firms may need to employ obfuscation strategies (Allender et al., 2021) and also account for strategic response from consumers.
- *Promotion* is a domain where personalized policies have long been implemented. We discuss the studies on the efficacy of personalization of promotional activities based on large-scale field experiments. Simester et al. (2020a) and Simester et al. (2020b) consider the personalization of promotions (e.g., discounted membership and 120-day free trials) to maximize profits, in the context of catalog marketer. Similarly, Yoganarasimhan et al. (2022) consider the personalization of the length of free trials for a software product to maximize subscriptions and revenue. One common finding in this literature is that personalization does not always work and that the efficacy of personalization depends on the exact method used. Policies based on lasso and XGBoost typically do well. However, policies based on many commonly

used outcome estimators, classification methods, and heterogeneous treatment effect estimators (e.g., regressions, k-nearest neighbors, random forests, causal tree, and generalized random forests) are shown to perform worse than a simple uniform policy based on Average Treatment Effect. This is because even the main effects of promotions are often small, which makes the signal-to-noise ratio needed for model learning in these cases high. Thus, a key takeaway is that personalizing promotions is not only tricky but may also lead to no significant improvement in profits.

6 Personalization and Welfare

As discussed in the previous section, personalization can sometimes lead to substantial gains for firms that employ such strategies. However, since firms optimize their outcome of interests, it is not clear whether personalization also improves consumer and social welfare. In principle, to the extent that the firm's outcome of interest is aligned with a welfare outcome, we can expect personalization to help increase that particular welfare outcome. However, in many cases, these two objectives are not aligned. For example, while personalized ad recommendation can help consumers find an ad that matches their interests, it comes at the expense of consumers' privacy. Further, the utilitarian analysis of welfare is complicated by the fact that these objectives (e.g., privacy and fairness) often have different units of measurement and are subjective by definition. As such, they are hard to measure and quantify. Therefore, research studies on the welfare implications of personalization often focus on a specific notion of welfare and view the problem through that specific lens.

In this section, we focus on four of these welfare notions that have been extensively studied in the literature: (1) search costs, (2) privacy, (3) fairness, and (4) polarization. We provide an overview of the research on how personalization interacts with each of these notions of welfare notions.

6.1 Search Cost

Reduction in consumers' search costs is a means through which personalization algorithms can enhance consumer welfare because better product recommendations make consumers spend less time searching for the options. Moreover, lab studies showed that personalization can shift consumers to search in a "choice mode", thereby resulting in an earlier stopping in the search process and lower search costs (Dellaert and Häubl, 2012). With the availability of search data, a series of studies further demonstrated the positive impact of personalized rankings on consumer welfare (Yoganarasimhan, 2020), and more broadly, the alignment between the firm's objective and consumer welfare (Ghose et al., 2012, 2014; Jeziorski and Segal, 2015; Donnelly et al., 2021). That is, personalized rankings derived by maximizing one objective (consumer welfare vs. firm's revenue) can improve the other objective. This alignment in the firm's objective and consumer welfare is an important piece in policy debates about personalization.

6.2 Privacy

While personalization algorithms help consumers find the right matches at a lower search cost, this increase in welfare often comes at the expense of consumer privacy. Personalization algorithms often require large-scale consumer-level data that are collected through tracking users online. The use of fine-grained user-level data, in turn, has heightened privacy concerns among consumers and privacy advocates. Prior research has documented these concerns in a variety of contexts through consumers' negative reaction to cases where the intrusiveness of targeting practices become more salient to them (Goldfarb and Tucker, 2011a; Acquisti et al., 2012), and their positive reaction when there is an increase in privacy controls and transparency within personalized policies (Tucker, 2014; Kim et al., 2019). Yet, there still remains a large gap between consumers' privacy concerns and actions, with only a tiny fraction of users opting out of behavioral targeting (Johnson et al., 2020). Nevertheless, a series of initiatives have led to significant limits to user tracking and behavioral targeting over the last few years. Prominent examples include GDPR by the European Union, which was enacted in 2018 (that requires firms to actively ask for permission before tracking users for advertising purposes, among other limits), and Apple's new iOS 14.5 that went into effect in 2021 (which blocks user tracking through device IDs in mobile devices). These initiatives have led to questions on the costs and benefits of preserving user privacy and who will bear the costs (if any) of these changes.

In particular, these measures have led to a renewed focus on the debate about the impact of privacy regulations on personalization and its value for Internet publishers. Goldfarb and Tucker (2011b) examined the impact of privacy regulations on Internet websites that monetize based on advertising and documented the negative of privacy regulations on these businesses. In particular, they show that this law most profoundly affects general content publishers such as new websites that rely more heavily on behavioral targeting. Rafieian and Yoganarasimhan (2021) study the interplay between personalization and privacy from the point-of-view of a platform that controls the level of targeting and focuses on the platform's incentives to protect consumers' privacy. They combine machine learning methods for personalization with the economic models of advertising auctions and show that too much targeting can soften the competition, thereby hurting the platform's revenues. This finding implies that the market mechanism creates some level of alignment between the platform's objective and that of consumers with respect to their privacy, thereby suggesting that self-regulation in the advertising marketplace may be possible.

Overall, one of the key takeaways from the literature on the interplay between personalization and privacy is that there is an inherent trade-off: privacy protection comes at the expense of the overall efficiency in the market. Finding the right balance is often challenging and requires a complete understanding of the marketplace and the incentives of different parties involved in this marketplace. Thus, it is essential to use a multi-method approach that combines the insights from market design literature with the AI machinery to understand the right balance between personalization and privacy from a welfare standpoint.

6.3 Fairness

Any deviation from a uniform policy for the entire population, by definition, begs the question of fairness: is it fair that different consumers receive different treatments? One of the earliest studies that brought this issue to attention was Sweeney (2013), who showed that ads for arrest records are more likely to show up in Google searches for black-sounding names compared to searches for white-sounding names. Anecdotes like these have led to significant interest and academic research on algorithmic fairness, which falls into three broad categories: (1) identification of disparities and discrimination, (2) potential explanation for such disparities, and (3) solutions to increase fairness in personalization.

The first stream of work focuses on identifying and quantifying the problem by defining context-relevant metrics (Barocas et al., 2019), developing algorithms to identify disparate impacts of personalized policies (Kallus and Zhou, 2019), and context-specific methods to audit for algorithmic discrimination (Imana et al., 2021). While the first stream of research identifies the problem, the second stream focuses on explaining why such disparities by examining the sources of the problem, such as data collection (Chen et al., 2018) or the market mechanism. For example, Lambrecht and Tucker (2019) studies the disparities in the likelihood of receiving STEM ads across men and women and identify the reason to be the higher cost of women's impressions in the advertising auctions. The third stream of this literature adopts an engineering lens to view the problem. It attempts to provide solutions for the problem through a variety of approaches, such as performing a constrained optimization subject to fairness constraints (Dwork et al., 2012), de-biasing data (Feldman et al., 2015), and only capturing heterogeneity that is unrelated to the protected classes in the learning stage (Ascarza and Israeli, 2021). Like privacy, fairness often comes at the expense of efficiency. The trade-off between efficiency and fairness has long been part of policy debates in many domains, some as unrelated to personalization as taxation. As such, an important area of research is to measure and quantify the price of fairness empirically (Bertsimas et al., 2011; Mehrotra et al., 2018; Kallus and Zhou, 2021b).

6.4 Polarization

The fact that the rise of political polarization in the US happened around the same time as the increase in personalized content delivery through online platforms has led to speculations about the causal impact of personalization on polarization (Pariser, 2011). The argument is that polarized content creates echo chambers where consumers only see the information that confirms their prior beliefs, thereby amplifying the political polarization. Early theoretical studies further suggested that personalization can contribute to polarization in a network context (Dandekar et al., 2013). These ideas and speculations motivated further research on the interaction between personalization and polarization with twin goals of: (1) empirically identifying the link in different contexts and (2) providing an algorithmic solution to the problem.

At a high level, empirical evidence does not paint a unified picture. The greatest increase in polarization in the US happened among demographic groups that are least likely to use the Internet (Boxell et al., 2017). Similarly, studies on users' browsing histories show that social media and the Internet increase users' exposure to opposing views, even though these channels are associated with greater ideological distance between users (Flaxman et al., 2016). Platform-level studies on the relationship between personalization and polarization also provide mixed evidence. Indeed, research on Facebook's news feed algorithm, Google's search personalization, and YouTube found little evidence for content partisanship that is attributable to personalization Bakshy et al. (2015); Robertson et al. (2018); Ribeiro et al. (2020); Hosseinmardi et al. (2021). Nevertheless, given the theoretical possibility of personalization leading to higher polarization and the growing anecdotal evidence on this topic, a growing body of work has focused on the re-design of personalization algorithms to actively control polarization (Celis et al., 2019), and guarantee the reachability of content (Dean et al., 2020).

7 Conclusion and Directions for Future Research

A fundamental goal shared by many businesses is to personalize all aspects of their products and interactions with users (e.g., marketing mix variables, user experiences). While this goal was largely aspirational in the past, the last decade saw significant advances in the practical feasibility of real-time and large-scale personalization. These advances have been made possible by two key factors: (1) the big leaps in computing power and data storage, which led to the development of powerful machine learning tools that can be used to personalize at scale, in real-time, and (2) the concurrent development of theoretical and statistical foundations of the algorithms and methods used for personalization.

This chapter reviewed the methods available for personalization, the correct way to evaluate personalized policies, examples of personalization across a series of contexts, and the welfare and policy issues that are often interwoven with personalization goals. That said, many difficult challenges remain, and effective personalization in the field remains elusive in many settings. We, therefore, conclude with a summary of these challenges and provide some directions for future research.

7.1 Signal-to-Noise Ratio

Even with all the advances in the machinery available, effective personalization and reliable evaluation of personalized policies remain a challenge *in practice*. One of the main reasons for this is the inherent noise in the data or outcomes, e.g., clicks and purchases are inherently stochastic. Prior researchers have hypothesized that this natural variability can create a "magic barrier" that may prevent us from both designing effective policies and getting accurate measures of the reward from a policy (Herlocker et al., 2004). For example, in an early paper Hill et al. (1995) showed that users provide inconsistent ratings when asked to rate the same movie at different times. They argue that an algorithm cannot be more accurate than the inherent variance in a user's ratings for the same item. This issue, also known as the low signal-to-noise problem in supervised learning, implies that the inherent noise in the data can make inference really challenging. As demonstrated in Yoganarasimhan et al. (2022), this can lead to situations where a personalized model can learn spurious patterns heterogeneity from the training data, which in turn can lead to an unreliable performance in a different data sample. Thus, future research will need to focus on robust methods immune to these noise problems.

7.2 Multiple Objectives and Long-term Outcomes

Typically, personalization models maximize an objective function that maximizes one outcome of interest (e.g., user engagement, revenue, clicks). However, in reality, most firms are interested in many different objectives, which may not all be aligned. For example, a social media firm may want to personalize a recommendation system that maximizes various consumer engagement metrics (e.g., posting new content, engaging with peers' content by liking the content, responding to older content). Further, these objectives can be differentially important for different customers. As of now, we lack good paradigms to build personalization models that can automatically accommodate multiple objectives.

A specific case of the multiple metrics problem is the short- vs. long-term metrics issue. In practice, many firms use short-term surrogates for personalization, even though their goal is often optimization of metrics based on a longer horizon. For example, Yoganarasimhan et al. (2022) consider a situation where the firm wants to personalize promotions (length of free trials). The outcome data available in the short-run for this task is subscriptions, even though the firm may care more about long-term metrics such as overall revenue or retention (which are only available after a year or two after the promotion). In principle, these two metrics can be highly misaligned. For example, a policy that increases subscriptions among students (who get a significant educational discount and hence pay lower prices) and/or users who subscribe to lower-end products/bundles (at low prices) at the expense of high-end users can lead to higher subscriptions but lower revenues. While they show that these objectives are largely aligned in their setting, these findings are not universally applicable. As such, we need more theory-driven models that allow us to draw a tighter link between long and short-term metrics. See Yang et al. (2021) for some recent developments in this area.

7.3 Time Drifts

Another challenge that is relevant in practical settings is time drift. For example, users' preferences can change with time, and/or the same actions/treatments can become stale over time. This can cause issues since almost all personalization models assume some constant response for any given action (for some user-level covariates). For example, typical bandit models assume that the rewards from a given arm are constant over time, which is often not true in practice. Sweeney et al. (2020) presents an excellent analysis of the effects of such model misspecifications for Bayesian bandits. Similarly, Simester et al. (2020b) provides some guidance on the supervised learning methods that are more robust to such drifts. Nevertheless, more research is necessary to develop algorithms robust to such drifts.

7.4 Strategic Behavior and Equilibrium Analysis

A series of early works on personalization and targeting formulated the problem in a strategic environment where the adoption and monetization of personalization are key decisions by the agents, without focusing on the details of the personalization algorithm (Shaffer and Zhang, 1995; Acquisti and Varian, 2005; Pancras and Sudhir, 2007; Zhang, 2011). On the other hand, much of the methodological literature on personalization focused on situations where a single firm personalizes actions in an environment where the other agents are non-strategic. This is unlikely to be true in most settings, especially on multi-sided platforms. A fruitful avenue for future research is to combine game-theoretical frameworks and personalization algorithms to generate strategy-proof policies.

In the context of advertising, Rafieian and Yoganarasimhan (2021) and Rafieian (2020) account for the presence of strategic bidding by advertisers when considering a platform's personalization decisions. For example, Rafieian and Yoganarasimhan (2021) shows that ad networks can benefit significantly from behavioral tracking and personalization if advertisers are non-strategic. However, if advertisers can strategically alter their bids, the ad network prefers lower levels of user-tracking and behavioral targeting. Thus, the platform's incentives are completely different, depending on whether the environment is strategic. In another work related to advertising auctions, Golrezaei et al. (2019) studies a dynamic case where a platform learns to set personalized reserve prices to optimize the auction revenues. However, they incorporate the fact that advertisers would have an incentive to shade their bids knowing that the platform would use them to set optimal reserve prices, thereby shrinking advertisers' surplus. Therefore, they focus on an incentive-aware algorithm that personalizes reserve prices while taking into account the strategic behavior of buyers.

We expect incentive-aware personalization that takes the equilibrium concept into account to be one of the most important areas of future research. Indeed, research on this topic has the potential to bring together ideas from all the main paradigms in marketing science – game theory, empirical structural models, machine learning, and causal inference – and provide answers to some of the most challenging marketing questions with broader policy implications.

References

A. Acquisti and H. R. Varian. Conditioning prices on purchase history. *Marketing Science*, 24(3): 367–381, 2005.

A. Acquisti, L. K. John, and G. Loewenstein. The impact of relative standards on the propensity to disclose. *Journal of Marketing Research*, 49(2):160–174, 2012.

A. Agarwal, S. Bird, M. Cozowicz, L. Hoang, J. Langford, S. Lee, J. Li, D. Melamed, G. Oshri, O. Ribas, et al. Making contextual decisions with low technical debt. *arXiv preprint arXiv:1606.03966*, 2016.

W. J. Allender, J. Liaukonyte, S. Nasser, and T. J. Richards. Price fairness and strategic obfuscation. *Marketing Science*, 40(1):122–146, 2021.

E. Angelino, M. J. Johnson, and R. P. Adams. Patterns of scalable bayesian inference. *arXiv preprint arXiv:1602.05221*, 2016.

A. Ansari and C. F. Mela. E-customization. *Journal of marketing research*, 40(2):131–145, 2003.

A. Ansari, S. Essegaier, and R. Kohli. Internet recommendation systems. *Journal of Marketing Research (JMR)*, 37(3), 2000.

A. Ansari, Y. Li, and J. Z. Zhang. Probabilistic topic model for hybrid recommender systems: A stochastic variational bayesian approach. *Marketing Science*, 37(6):987–1008, 2018.

E. Ascarza. Retention futility: Targeting high-risk customers might be ineffective. *Journal of Marketing Research*, 55(1):80–98, 2018.

E. Ascarza and A. Israeli. Beat unintended bias in personalized policies. *Available at SSRN 3908088*, 2021.

S. Athey and G. Imbens. Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, 113(27):7353–7360, 2016.

S. Athey and S. Wager. Policy learning with observational data. *Econometrica*, 89(1):133–161, 2021.

S. Athey, J. Tibshirani, S. Wager, et al. Generalized random forests. *The Annals of Statistics*, 47(2): 1148–1178, 2019.

P. Auer. Using confidence bounds for exploitation-exploration trade-offs. *Journal of Machine Learning Research*, 3(Nov):397–422, 2002.

E. Bakshy, S. Messing, and L. A. Adamic. Exposure to ideologically diverse news and opinion on facebook. *Science*, 348(6239):1130–1132, 2015.

S. Barocas, M. Hardt, and A. Narayanan. *Fairness and Machine Learning*. fairmlbook.org, 2019. `http://www.fairmlbook.org`.

H. Bastani and M. Bayati. Online decision making with high-dimensional covariates. *Operations Research*, 68(1):276–294, 2020.

H. Bastani, M. Bayati, and K. Khosravi. Mostly exploration-free algorithms for contextual bandits. *Management Science*, 67(3):1329–1349, 2021a.

H. Bastani, P. Harsha, G. Perakis, and D. Singhvi. Learning personalized product recommendations with customer disengagement. *Manufacturing & Service Operations Management*, 2021b.

D. Bertsimas, V. F. Farias, and N. Trichakis. The price of fairness. *Operations research*, 59(1): 17–31, 2011.

A. Bietti, A. Agarwal, and J. Langford. A contextual bandit bake-off. *Journal of Machine Learning Research*, 22(133):1–49, 2021.

D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.

L. Boxell, M. Gentzkow, and J. M. Shapiro. Greater internet use is not associated with faster growth in political polarization among us demographic groups. *Proceedings of the National Academy of Sciences*, 114(40):10612–10617, 2017.

E. J. Candès and B. Recht. Exact matrix completion via convex optimization. *Foundations of Computational mathematics*, 9(6):717–772, 2009.

E. J. Candès and T. Tao. The power of convex relaxation: Near-optimal matrix completion. *IEEE Transactions on Information Theory*, 56(5):2053–2080, 2010.

A. Capponi, S. Olafsson, and T. Zariphopoulou. Personalized robo-advising: Enhancing investment through client interaction. *Management Science*, 2021.

L. E. Celis, S. Kapoor, F. Salehi, and N. Vishnoi. Controlling polarization in personalization: An algorithmic framework. In *Proceedings of the conference on fairness, accountability, and transparency*, pages 160–169, 2019.

O. Chapelle, E. Manavoglu, and R. Rosales. Simple and Scalable Response Prediction for Display Advertising. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 5(4):61, 2015.

I. Chen, F. D. Johansson, and D. Sontag. Why is my classifier discriminatory? *Advances in Neural Information Processing Systems*, 31, 2018.

V. Chernozhukov, J. C. Escanciano, H. Ichimura, W. K. Newey, and J. M. Robins. Locally robust semiparametric estimation. *arXiv preprint arXiv:1608.00033*, 2016.

V. Chernozhukov, M. Demirer, E. Duflo, and I. Fernandez-Val. Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in india. Technical report, National Bureau of Economic Research, 2018.

J. Chung and V. R. Rao. A general consumer preference model for experience products: application to internet recommendation services. *Journal of marketing research*, 49(3):289–305, 2012.

T. S. Chung, R. T. Rust, and M. Wedel. My mobile music: An adaptive personalization system for digital audio players. *Marketing Science*, 28(1):52–68, 2009.

A. Coenen. How The New York Times is Experimenting with Recommendation Algorithms, 2019. URL `https://open.nytimes.com/how-the-new-york-times-is-experimenting-with-recommendation-algorithms-56`

P. Dandekar, A. Goel, and D. T. Lee. Biased assimilation, homophily, and the dynamics of polarization. *Proceedings of the National Academy of Sciences*, 110(15):5791–5796, 2013.

S. Dean, S. Rich, and B. Recht. Recommendations and user agency: the reachability of collaboratively-filtered information. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pages 436–445, 2020.

B. G. Dellaert and G. Häubl. Searching in choice mode: Consumer decision processes in product search with recommendations. *Journal of Marketing Research*, 49(2):277–288, 2012.

R. Donnelly, A. Kanodia, and I. Morozov. Welfare effects of personalized rankings. *Available at SSRN 3649342*, 2021.

J.-P. Dubé and S. Misra. Personalized pricing and customer welfare. 2021.

M. Dudík, J. Langford, and L. Li. Doubly robust policy evaluation and learning. In *Proceedings of the 28th International Conference on International Conference on Machine Learning*, pages 1097–1104. Omnipress, 2011.

C. Dwork, M. Hardt, T. Pitassi, O. Reingold, and R. Zemel. Fairness through awareness. In *Proceedings of the 3rd innovations in theoretical computer science conference*, pages 214–226, 2012.

A. Farahat and M. C. Bailey. How effective is targeted advertising? In *Proceedings of the 21st international conference on World Wide Web*, pages 111–120, 2012.

V. F. Farias and A. A. Li. Learning preferences with side information. *Management Science*, 65(7): 3131–3149, 2019.

M. Feldman, S. A. Friedler, J. Moeller, C. Scheidegger, and S. Venkatasubramanian. Certifying and removing disparate impact. In *proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, pages 259–268, 2015.

S. Flaxman, S. Goel, and J. M. Rao. Filter bubbles, echo chambers, and online news consumption. *Public opinion quarterly*, 80(S1):298–320, 2016.

D. Foster and A. Rakhlin. Beyond ucb: Optimal and efficient contextual bandits with regression oracles. In *International Conference on Machine Learning*, pages 3199–3210. PMLR, 2020.

A. Ghose, P. G. Ipeirotis, and B. Li. Designing ranking systems for hotels on travel search engines by mining user-generated and crowdsourced content. *Marketing Science*, 31(3):493–520, 2012.

A. Ghose, P. G. Ipeirotis, and B. Li. Examining the impact of ranking on consumer behavior and search engine revenue. *Management Science*, 60(7):1632–1654, 2014.

D. Goldberg, D. Nichols, B. M. Oki, and D. Terry. Using collaborative filtering to weave an information tapestry. *Communications of the ACM*, 35(12):61–70, 1992.

A. Goldfarb and C. Tucker. Online Display Advertising: Targeting and Obtrusiveness. *Marketing Science*, 30(3):389–404, 2011a.

A. Goldfarb and C. E. Tucker. Privacy Regulation and Online Advertising. *Management science*, 57(1):57–71, 2011b.

A. Goli, D. Reiley, and H. Zhang. Personalized versioning: Product strategies constructed from experiments on pandora. Working Paper, 2021.

A. Goli, A. Lambrecht, and H. Yoganarasimhan. A bias correction approach for interference in ranking experiments. Working Paper, 2022.

N. Golrezaei, A. Javanmard, and V. Mirrokni. Dynamic incentive-aware learning: Robust pricing in contextual auctions. *Advances in Neural Information Processing Systems*, 32, 2019.

C. A. Gomez-Uribe and N. Hunt. The netflix recommender system: Algorithms, business value, and innovation. *ACM Transactions on Management Information Systems (TMIS)*, 6(4):1–19, 2015.

L. Guelman, M. Guillén, and A. M. Pérez-Marín. Uplift random forests. *Cybernetics and Systems*, 46(3-4):230–248, 2015.

J. R. Hauser, G. L. Urban, G. Liberali, and M. Braun. Website morphing. *Marketing Science*, 28(2): 202–223, 2009.

X. He, J. Pan, O. Jin, T. Xu, B. Liu, T. Xu, Y. Shi, A. Atallah, R. Herbrich, S. Bowers, et al. Practical Lessons from Predicting Clicks on Ads at Facebook. In *Proceedings of the Eighth International Workshop on Data Mining for Online Advertising*, pages 1–9. ACM, 2014.

J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl. Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems (TOIS)*, 22(1):5–53, 2004.

J. L. Hill. Bayesian nonparametric modeling for causal inference. *Journal of Computational and Graphical Statistics*, 20(1):217–240, 2011.

W. Hill, L. Stead, M. Rosenstein, and G. Furnas. Recommending and evaluating choices in a virtual community of use. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 194–201, 1995.

G. J. Hitsch and S. Misra. Heterogeneous treatment effects and optimal targeting policy evaluation. *Available at SSRN 3111957*, 2018.

D. G. Horvitz and D. J. Thompson. A generalization of sampling without replacement from a finite universe. *Journal of the American statistical Association*, 47(260):663–685, 1952.

H. Hosseinmardi, A. Ghasemian, A. Clauset, M. Mobius, D. M. Rothschild, and D. J. Watts. Examining the consumption of radical content on youtube. *Proceedings of the National Academy of Sciences*, 118(32), 2021.

B. Imana, A. Korolova, and J. Heidemann. Auditing for discrimination in algorithms delivering job ads. In *Proceedings of the Web Conference 2021*, pages 3767–3778, 2021.

B. Jacobs, D. Fok, and B. Donkers. Understanding large-scale dynamic purchase behavior. *Marketing Science*, 40(5):844–870, 2021.

B. J. Jacobs, B. Donkers, and D. Fok. Model-based purchase predictions for large assortments. *Marketing Science*, 35(3):389–404, 2016.

P. Jeziorski and I. Segal. What makes them click: Empirical analysis of consumer demand for search advertising. *American Economic Journal: Microeconomics*, 7(3):24–53, 2015.

N. Jiang and L. Li. Doubly robust off-policy value evaluation for reinforcement learning. In *International Conference on Machine Learning*, pages 652–661. PMLR, 2016.

Z. Jiang, T. Chan, H. Che, and Y. Wang. Consumer search and purchase: An empirical investigation of retargeting based on consumer online behaviors. *Marketing Science*, 40(2):219–240, 2021.

G. A. Johnson, S. K. Shriver, and S. Du. Consumer privacy choice in online advertising: Who opts out and at what cost to industry? *Marketing Science*, 39(1):33–51, 2020.

N. Kallus and M. Udell. Dynamic assortment personalization in high dimensions. *Operations Research*, 68(4):1020–1037, 2020.

N. Kallus and M. Uehara. Double reinforcement learning for efficient off-policy evaluation in markov decision processes. *J. Mach. Learn. Res.*, 21:167–1, 2020.

N. Kallus and A. Zhou. Assessing disparate impact of personalized interventions: identifiability and bounds. *Advances in Neural Information Processing Systems*, 32, 2019.

N. Kallus and A. Zhou. Minimax-optimal policy learning under unobserved confounding. *Management Science*, 67(5):2870–2890, 2021a.

N. Kallus and A. Zhou. Fairness, welfare, and equity in personalized pricing. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pages 296–314, 2021b.

C. Kasula. Netflix recommender system – a big data case study, 2020. URL `https://towardsdatascience.com/netflix-recommender-system-a-big-data-case-study-19cfa6d56ff5`.

T. Kim, K. Barasz, and L. K. John. Why am i seeing this ad? the effect of ad transparency on ad effectiveness. *Journal of Consumer Research*, 45(5):906–932, 2019.

T. Kitagawa and A. Tetenov. Who should be treated? empirical welfare maximization methods for treatment choice. *Econometrica*, 86(2):591–616, 2018.

R. Kohavi, D. Tang, and Y. Xu. *Trustworthy online controlled experiments: A practical guide to a/b testing*. Cambridge University Press, 2020.

A. Lambrecht and C. Tucker. Algorithmic bias? an empirical study of apparent gender-based discrimination in the display of stem career ads. *Management science*, 65(7):2966–2981, 2019.

J. Langford and T. Zhang. The epoch-greedy algorithm for multi-armed bandits with side information. *Advances in neural information processing systems*, 20, 2007.

L. Li, W. Chu, J. Langford, and R. E. Schapire. A contextual-bandit approach to personalized news article recommendation. In *Proceedings of the 19th international conference on World wide web*, pages 661–670, 2010.

S. Li, J. Kawale, and Y. Fu. Deep collaborative filtering via marginalized denoising auto-encoder. In *Proceedings of the 24th ACM international on conference on information and knowledge management*, pages 811–820, 2015.

P. Liao, P. Klasnja, and S. Murphy. Off-policy estimation of long-term average outcomes with applications to mobile health. *Journal of the American Statistical Association*, 116(533):382–391, 2021.

J. Liaukonyte. Personalized and social commerce. Working Paper, 2021.

G. Liberali and A. Ferecatu. Morphing for consumer dynamics: Bandits meet hidden markov models. *Marketing Science*, 2022.

G. Linden, B. Smith, and J. York. Amazon. com recommendations: Item-to-item collaborative filtering. *IEEE Internet computing*, 7(1):76–80, 2003.

T. Mandel, Y.-E. Liu, S. Levine, E. Brunskill, and Z. Popovic. Offline policy evaluation across representations with applications to educational games. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, pages 1077–1084. International Foundation for Autonomous Agents and Multiagent Systems, 2014.

H. B. McMahan, G. Holt, D. Sculley, M. Young, D. Ebner, J. Grady, L. Nie, T. Phillips, E. Davydov, D. Golovin, et al. Ad Click Prediction: A View from the Trenches. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1222–1230. ACM, 2013.

R. Mehrotra, J. McInerney, H. Bouchard, M. Lalmas, and F. Diaz. Towards a fair marketplace: Counterfactual evaluation of the trade-off between relevance, fairness & satisfaction in recommendation systems. In *Proceedings of the 27th acm international conference on information and knowledge management*, pages 2243–2251, 2018.

A. L. Montgomery, S. Li, K. Srinivasan, and J. C. Liechty. Modeling online browsing and path analysis using clickstream data. *Marketing science*, 23(4):579–595, 2004.

I. Morozov, S. Seiler, X. Dong, and L. Hou. Estimation of preference heterogeneity in markets with costly search. *Marketing Science*, 40(5):871–899, 2021.

B. Murthi and S. Sarkar. The role of the management sciences in research on personalization. *Management Science*, 49(10):1344–1362, 2003.

X. Nie and S. Wager. Quasi-oracle estimation of heterogeneous treatment effects. *Biometrika*, 108 (2):299–319, 2021.

J. Pancras and K. Sudhir. Optimal marketing strategies for a customer data intermediary. *Journal of Marketing research*, 44(4):560–578, 2007.

E. Pariser. *The filter bubble: How the new personalized web is changing what we read and how we think*. Penguin, 2011.

D. Proserpio, J. R. Hauser, X. Liu, T. Amano, A. Burnap, T. Guo, D. D. Lee, R. Lewis, K. Misra, E. Schwarz, et al. Soul and machine (learning). *Marketing Letters*, 31(4):393–404, 2020.

O. Rafieian. Revenue-optimal dynamic auctions for adaptive ad sequencing. Technical report, Working paper, 2020.

O. Rafieian. Optimizing user engagement through adaptive ad sequencing. Technical report, Working paper, 2022.

O. Rafieian and H. Yoganarasimhan. Targeting and privacy in mobile advertising. *Marketing Science*, 2021.

B. Recht. A simpler approach to matrix completion. *Journal of Machine Learning Research*, 12 (12), 2011.

M. H. Ribeiro, R. Ottoni, R. West, V. A. Almeida, and W. Meira Jr. Auditing radicalization pathways on youtube. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*, pages 131–141, 2020.

R. E. Robertson, S. Jiang, K. Joseph, L. Friedland, D. Lazer, and C. Wilson. Auditing partisan audience bias within google search. *Proceedings of the ACM on Human-Computer Interaction*, 2 (CSCW):1–22, 2018.

P. E. Rossi, R. E. McCulloch, and G. M. Allenby. The value of purchase history data in target marketing. *Marketing Science*, 15(4):321–340, 1996.

P. Rzepakowski and S. Jaroszewicz. Decision trees for uplift modeling with single and multiple treatments. *Knowledge and Information Systems*, 32(2):303–327, 2012.

B. Sarwar, G. Karypis, J. Konstan, and J. Riedl. Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on World Wide Web*, pages 285–295, 2001.

G. Shaffer and Z. J. Zhang. Competitive coupon targeting. *Marketing Science*, 14(4):395–416, 1995.

U. Shalit, F. D. Johansson, and D. Sontag. Estimating individual treatment effect: generalization bounds and algorithms. In *International Conference on Machine Learning*, pages 3076–3085. PMLR, 2017.

D. Simchi-Levi and Y. Xu. Bypassing the monster: A faster and simpler optimal algorithm for contextual bandits under realizability. *Mathematics of Operations Research*, 2021.

D. Simester, A. Timoshenko, and S. I. Zoumpoulis. Efficiently evaluating targeting policies: Improving on champion vs. challenger experiments. *Management Science*, 66(8):3412–3424, 2020a.

D. Simester, A. Timoshenko, and S. I. Zoumpoulis. Targeting prospective customers: Robustness of machine-learning methods to typical data challenges. *Management Science*, 66(6):2495–2522, 2020b.

A. Singhal, P. Sinha, and R. Pant. Use of deep learning in modern recommendation system: A summary of recent works. *arXiv preprint arXiv:1712.07525*, 2017.

J. H. Steckel, R. S. Winer, R. E. Bucklin, B. G. Dellaert, X. Drèze, G. Häubl, S. D. Jap, J. D. Little, T. Meyvis, A. L. Montgomery, et al. Choice in interactive environments. *Marketing Letters*, 16 (3):309–320, 2005.

A. Swaminathan and T. Joachims. Counterfactual risk minimization: Learning from logged bandit feedback. In *International Conference on Machine Learning*, pages 814–823. PMLR, 2015.

L. Sweeney. Discrimination in online ad delivery: Google ads, black names and white names, racial discrimination, and click advertising. *Queue*, 11(3):10–29, 2013.

M. Sweeney, M. van Adelsberg, K. Laskey, and C. Domeniconi. Effects of model misspecification on bayesian bandits case studies in ux optimization. In *2020 IEEE International Conference on Data Mining (ICDM)*, pages 1286–1291. IEEE, 2020.

A. Tewari and S. A. Murphy. From ads to interventions: Contextual bandits in mobile health. In *Mobile Health*, pages 495–517. Springer, 2017.

G. Theocharous, P. S. Thomas, and M. Ghavamzadeh. Personalized ad recommendation systems for life-time value optimization with guarantees. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2015.

P. Thomas and E. Brunskill. Data-efficient off-policy policy evaluation for reinforcement learning. In *International Conference on Machine Learning*, pages 2139–2148. PMLR, 2016.

C. E. Tucker. Social Networks, Personalized Advertising, and Privacy Controls. *Journal of Marketing Research*, 51(5):546–562, 2014.

G. L. Urban, G. Liberali, E. MacDonald, R. Bordley, and J. R. Hauser. Morphing banner advertising. *Marketing Science*, 33(1):27–46, 2014.

S. Wager and S. Athey. Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 0(0):1–15, 2018. doi: 10.1080/01621459.2017.1319839.

X. Wang, L. Yu, K. Ren, G. Tao, W. Zhang, Y. Yu, and J. Wang. Dynamic attention deep model for article recommendation by learning human editors' demonstration. In *Proceedings of the 23rd acm sigkdd international conference on knowledge discovery and data mining*, pages 2051–2059, 2017a.

Y.-X. Wang, A. Agarwal, and M. Dudık. Optimal and adaptive off-policy evaluation in contextual bandits. In *International Conference on Machine Learning*, pages 3589–3597. PMLR, 2017b.

J. Yang, D. Eckles, P. Dhillon, and S. Aral. Targeting for long-term outcomes. *arXiv preprint arXiv:2010.15835*, 2021.

S. Yao and C. F. Mela. A dynamic model of sponsored search advertising. *Marketing Science*, 30 (3):447–468, 2011.

Y. Ying, F. Feinberg, and M. Wedel. Leveraging missing ratings to improve online recommendation systems. *Journal of marketing research*, 43(3):355–365, 2006.

H. Yoganarasimhan. Search personalization using machine learning. *Management Science*, 66(3): 1045—-1070, 2020.

H. Yoganarasimhan, E. Barzegary, and A. Pani. Design and evaluation of optimal free trials. *Management Science*, 2022.

J. Zhang. The perils of behavior-based personalization. *Marketing Science*, 30(1):170–186, 2011.

J. Zhang and L. Krishnamurthi. Customizing promotions in online stores. *Marketing science*, 23(4): 561–578, 2004.